Approved for Public Release;
Distribution Unlimited. 14-3513

MP140633
October 17, 2014

Julie McEwen
Stuart Shapiro
Julie Snyder
Duane Blackburn

MITRE Response to OSTP/NITRD "National Privacy Research Strategy" RFI

For additional information about this response, please contact:

Duane Blackburn
dblackburn@mitre.org
(434) 964-5023

This page intentionally left blank.

Table of Contents

Introduction

The MITRE Corporation is a not for profit company that runs Federally-Funded Research and Development Centers (FFRDCs) for the U.S. government. MITRE's FFRDCs serve agencies in a variety of areas that impact the public in direct and indirect ways, such as national security; aviation safety and administration; tax administration; homeland security; healthcare; benefits services; cybersecurity; and other missions. We are pleased to respond to your RFI regarding a National Privacy Research Strategy based on our broad perspective gained from serving a variety of government missions, and from the unique perspective of a systems engineering company that combines a strong research base with an informed awareness of the larger policy and contexts in which government operations are conducted.

Protecting the privacy of the data about individuals that government systems contain is a critical pillar of our sponsors' information technology programs. The public needs to trust that the government is keeping personal data safe from misuse. These same needs also apply to private-sector entities, who often hold much more data (and much more sensitive data) than government entities, but often aren't as knowledgeable about the risks of that data or the best ways to protect it. Finally, recent "hacks" of deeply personal photos of celebrities highlight one of the greatest needs in a privacy research strategy: developing methods that will enable individuals to understand their privacy risks across multiple domains, and how they can best protect themselves.

MITRE's privacy approach emphasizes strategy and policy as well as privacy engineering, and spans all aspects of a privacy program. It's about more than just complying with the law. Our broad view of privacy begins with the concept of a framework that includes technical, operational, social, and ethical implications of designs and processes. We help government agencies to:

- Comply with the letter and spirit of privacy laws and regulations
- Build trust and respect among constituents
- Facilitate appropriate sharing of personally identifiable information (PII)
- Reduce threats to personally identifiable information, such as identity theft and insider threats
- Align their privacy policies with mission objectives
- Plan and execute their privacy programs strategically
- Systematically build privacy into systems ("privacy engineering")

It is through this background that we provide our response to this RFI.

As individuals, elements of our world become more connected every day. The abilities to compartmentalize certain elements of our lives (e.g., professional and social personas), control what information is tracked and shared (e.g., in-store shopping habits), and even control what data is left behind anytime we use a device that is part of the rapidly expanding Internet of Things (e.g., "digital exhaust") are diminishing at a pace that, so far, greatly exceeds the ability of law and scholars to determine how to harness the desirable traits of these technologies in a way that does not erode the notions of privacy that we hold dear. The massive amounts of data we unwittingly create about ourselves results in the reality where organizations know us better than we may want them to, and even perhaps better than we know ourselves. There is a near-constant tension between utility in technologies and some of the many notions that comprise "privacy". There are myriad examples of each of the "abilities" people are losing that impacts privacy.

In 2007, one of the key news stories that illuminated the clashes social media can raise between professional and social personas was that of a student that was denied a teaching degree based on a photo posted on her MySpace account[1]. Between now and then, there are many other examples of social media posts resulting in employment sanctions, including workers that were fired. Without clear boundaries, some employers started demanding access to social media login credentials for job-seekers as a condition of employment, followed by a number of states passing laws against such activities. As stated by the National Conference of State Legislatures (NCSL), "Some employers argue that access to personal accounts is needed to protect proprietary information or trade secrets, to comply with federal financial regulations, or to prevent the employer from being exposed to legal liabilities. But others consider requiring access to personal accounts an invasion of employee privacy."[2] The description of "privacy" is vague in that statement, but it is likened to the First Amendment (specifically, freedom of speech) and a general ability to freely act as oneself.

Worker privacy continues to be a topic of debate today, but it is only one of the ongoing privacy discussions related to social media. Social media sites are also a source of privacy breaches, such as the stolen images that are later used to create fake profiles to conduct malicious activities, phishing scams that entice users to click on links that lead to malware, and in other potentially compromising ways (e.g., advertising illegal services). Social media companies, such as Facebook, are expanding their privacy offerings and technical capabilities to manage access, but the frequent changes can pose an administrative burden and create confusion within their user communities.

Consumer shopping and purchasing habits are another area where privacy concerns are on the rise. Click streams and ad tracking technologies have been an integral part of the online experience for some time, especially for retail sites. Thanks in part to consumer concerns and the heightened awareness of privacy issues on the Internet in recent years, new tools and browsers are cropping up with promises of protecting user privacy. A newer trend that is garnering attention is the use of mobile devices to track consumer shopping habits in brick and mortar stores. Technology now allows retailers to use device identifiers to track consumers as they move throughout the store, and it can be combined with video feeds from "security" cameras which are now used for more than addressing security incidents (e.g., -tracking shopper traffic).

The age of connecting to the Internet solely through a computer is long gone. Now everything from phones to gaming consoles to smart meters to watches to TVs to home appliances to home security devices to cars can connect through WiFi. According to a

[1] http://www.foxnews.com/story/2007/04/29/would-be-teacher-denied-degree-over-drunken-pirate-myspace-photo-sues/

[2] http://www.ncsl.org/research/telecommunications-and-information-technology/employer-access-to-social-media-passwords-2013.aspx. This page also tracks legislative activity regarding this topic in the states.

February 2014 study for TRUSTe[3], 59% of Internet users know that smart devices can collect information about their personal activities and only 22% of those surveyed felt the benefit of smart devices out-weighed any privacy concerns. Yet new devices and technologies throughout multiple industry sectors get "smarter" and more connected every day. Currently, it appears the only thing that may slow this evolution is the amount of bandwidth available to support these devices. It appears the pace of technology development, along with the associated unprecedented increase in data produced, is far out-pacing the breadth of consumer understanding. This disconnect between the rapid growth of device connectivity and the much slower growth in consumer awareness and understanding points to how wide the privacy divide is in this area and the need for privacy integration into the development of these products.

Further complicating all of the data that is amassed is the concept of the "right to be forgotten," which is now law in the European Union and a concept consumers elsewhere are likely to expect in certain scenarios. In this area, the privacy considerations collide with other core societal values, such as freedom of speech and censorship on the Internet. The Internet effectively has a permanent memory, rendering this concept difficult to apply in practice, particularly for search engines and the like.

Each trend discussed above raises interesting privacy challenges. The totality of these, and other, trends in light of the speed of evolution puts the nation on a trajectory where the "is privacy dead?" question that is frequently debated would be answered with a resounding "Yes!" if action is not taken. This clearly points to a need for privacy considerations to be integrated into product development in their earliest stages. While doing so, it must be clear that *privacy* is not synonymous with *security*.

While privacy focuses on the individual's ability to control the collection, use, and dissemination of their PII, security provides the mechanisms to ensure confidentiality and integrity of information, and the availability of information technology systems. The concepts of privacy and security, however, do intersect. Specifically, certain IT controls established to ensure confidentiality and integrity from a security perspective also support privacy objectives. For example, access controls ensure that only authorized individuals can read, alter, or delete PII. Such controls help achieve confidentiality and integrity from a security standpoint. In addition, when a system processes or stores PII, these IT controls ensure that users can access only the specific PII needed to perform their jobs; this helps ensure that use of PII is limited to authorized purposes (*purpose specification*) and protected from unauthorized access, destruction, and disclosure (*security safeguards*). While establishing good security practices helps protect privacy, these practices are not, in and of themselves, sufficient to fully address the Fair Information Practice Principles (FIPPs).

Many organizations rely on the following activities to address privacy risks:

- Policy
- Risk assessments (e.g., PIAs)
- Notice

[3] Results of the survey are discussed here: http://www.truste.com/events/iot/2014/05/59-of-u-s-internet-users-know-smart-devices-can-collect-information-about-their-personal-activities/. An infographic summary of the results are available here: http://www.truste.com/us-internet-of-things-index-2014/.

- Records management
- Accounting of disclosures
- Data flow mapping
- Data loss prevention
- Metrics

Yet privacy risks remain and privacy breaches continue to rise. Why? Because these things alone do not *proactively a*ddress privacy risks at the appropriate level of specificity for a given system. To be effective, systems containing PII must be capable of preventing or minimizing the effect of human error or fallibility and appropriately constraining system actions.

To adequately address privacy risks, systems that manage PII must behave in a privacy-sensitive manner. Systems engineering processes are a largely untapped opportunity to embed privacy requirements into organizational activities in a way that provides major impact and will proactively address privacy risks.

However, in examining how to include privacy in systems engineering processes, a holistic approach needs to be taken. The focus should be on socio-technical systems or systems more broadly construed with the objective of having both the people aspects and technical aspects of organizations work together to meet privacy requirements. This necessitates linking people aspects and business operations with the more technical aspects of systems engineering efforts.

The sections below provide more insight into the privacy objectives as viewed from current scenarios where privacy is a concern, risk assessment methods, multi-disciplinary approaches, and privacy architectures that should all be addressed in a privacy research strategy.

Question 1: Privacy objectives

Describe one or more scenarios that illustrate a critical issue concerning privacy; describe what privacy problems arise in the scenario; describe why it is important to overcome the identified problems; describe the needed privacy and what capabilities are required to achieve it; and describe what barriers exist to achieving the needed privacy in the scenario. The use of particular domains in the scenario (e.g., healthcare, education, social media) to describe the desired privacy state is encouraged.

Healthcare represents an area where rapid changes are occurring to the entire industry. Care providers are adopting electronic health records (EHRs) and electronic prescribing capabilities in the wake of the Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009. The resulting increase in electronic health data provides a rich store of information upon which to potentially conduct data analytics. Now, when many consumers visit a pharmacy to pick up a prescription, their doctor has conveniently sent their prescription electronically. The consumer may pay with a credit card, and may even allow the pharmacy to scan their store loyalty card. Multiple systems are likely used to conduct all of these activities. The pharmacy may pool information from those systems into a decision support system and perform data analytics in order to understand how it is tracking against quality and safety measures, prescription adherence for refills, and other metrics, as well as its customers buying habits.

In the rush to receive incentives by demonstrating meaningful use, organizations may not have configured all of the appropriate policies, procedures, and reporting templates to prohibit the unauthorized disclosure of PII. Perhaps even worse, organizations may not have thought about the bigger impacts of combining so much intensely personal information. Pharmacies are likely to know what kinds of illnesses people have (based on prescription and possibly on-site clinics) and, depending on buying habits, may even have a clue as to why the individual has the illness. For example, if a diabetic person periodically stocks up on candy, they may aggravate their symptoms. They may have a pattern that shows periodic candy purchases, followed by an increase in insulin ordered. Consumers have limited visibility into what information is tracked and how it is used, and limited control over the information about them in their loyalty reward accounts.

Prescription adherence metrics is a particularly interesting area. Prescription adherence metrics are limited in the amount of verifiably accurate information they can indicate, such as whether a particular prescription was refilled within the expected number of days, leading to assumptions about whether the consumer is taking the medication as indicated by their doctor. However, it is conceivable that pharmacies may make assumptions around this metric and begin to market services related to that. For example, pharmacies have the information available to perform the following actions:

- Inform doctors that a consumer is possibly not taking as much of their medication as prescribed,
- Inform insurance companies that medication is potentially going unused and that the consumer may not be following "doctor's orders"
- Increase marketing to the consumer regarding prescription reminders, alternatives to their prescriptions, products that combat known side effects of their prescriptions, or other products and services related to their assumed condition
- Trigger more consumer counseling with the pharmacist
- Refuse to fill prescriptions for a consumer that appears to regularly fail to comply with prescription orders

There are legitimate healthcare reasons to track prescription adherence, particularly when it comes to treating chronic conditions. For example, research has shown that compared with patients who follow instructions, patients who don't take their medications as intended have a risk for hospitalization, rehospitalization, and premature death that is 5.4 times higher if they have hypertension, 2.8 times higher if they have dyslipidemia, and 1.5 times higher if they have heart disease.[4] However, patients don't take their medications for a multitude of reasons which are not fully understood both by patients and medical researchers[5] In 2009, a team of researchers at Kaiser Permanente combed through much of the vast literature on compliance and distilled the sea of data down to several important patient-related barriers.[6] They include forgetfulness; lack of knowledge about the medication and its use; cultural, health, and/or religious beliefs about the medication;

[4] Claxton AJ, Cramer J, Pierce C. A systematic review of the associations between dose regimens and medication compliance. Clin Ther. 2012;23:1296-1310.

[5] Why Are So Many Patients Noncompliant?, Neil Chesanow, January 16, 2014, http://www.medscape.com/viewarticle/818850

[6] Oyekan E, Nimalasuriya A, Martin J, et al. The B-SMART appropriate medication-use process: a guide for clinicians to help patients -- part 1: barriers, solutions, and motivation. Permanente J. 2009;13:62-69.

denial or ambivalence regarding the state of their health; financial challenges; lack of health literacy; and lack of social support.[7] Other factors, which may or may not relate to the patient's health, could influence adherence, such as side effects, desire for natural alternatives, or a patient decision to end treatment.

Given the complexities that surround consumer decisions regarding prescription adherence, making assumptions and especially taking actions based on those assumptions could be contentious for all involved. Assuming a pharmacy were planning to implement the possible actions related to prescription adherence, tracking notice and consent of each consumer would be a critical element of the program, as would active consumer participation. One key consideration is how much effort it takes the consumer to maintain their desired level of privacy. Managing through this effectively and with limited burden to the consumer will require the support of technology and thoughtful privacy engineering practices. Consumers would likely need greater access to view the information maintained about them and to express their preferences for participating in such a program, marketing, sharing, and other activities. Data would need to be properly tagged using a consistent tagging scheme that all participating systems could interpret and handle accordingly.

System architecture and design would need to support all of these capabilities, many of which will rely on technologies to achieve. However, technology alone does not necessarily achieve all privacy objectives, or address all privacy risks. Identifying privacy requirements such as those described above should be part of a larger, systemic process to examine privacy risks, identify privacy requirements, and verify that system design supports those requirements. Collectively, these activities are referred to as "privacy engineering".

Question 2: Assessment capabilities

Discuss concepts, methods, and constructs needed to assess privacy; discuss capabilities and models that can: Express privacy requirements, assess and quantify risks/benefits to privacy, evaluate effects of privacy risk mitigation, and determine the fulfillment of privacy requirements.

The Fair Information Practice Principles (FIPPs) are a generally-recognized basis for information privacy in both the public and private sectors around the world. This includes various widely referenced, very similar versions of FIPPs, such as the Organization for Economic Cooperation and Development (OECD) Privacy Framework (previously the Guidelines on the Protection of Privacy and Transborder Flows of Personal Data)[8], the Asia-Pacific Economic Cooperation (APEC) Privacy Framework, and the Generally Accepted Privacy Principles (GAPP) jointly developed by the American Institute of CPAs and the Canadian Institute of Chartered Accountants. This commonality is reflected in the definition of FIPPs found in Appendix A of the National Strategy for Trusted Identities in

[7] Ibid 5.

[8] Appendix A of this paper lists the websites where these documents can be found.

Cyberspace (NSTIC)[9], and the glossary of the International Association of Privacy Professionals (IAPP).[10]

There are also a number of substantive methodologies grounded in the FIPPs, such as the Privacy Management Reference Model and Methodology (PMRM) developed by the Organization for the Advancement of Structured Information Standards (OASIS), the Privacy Evaluation Methodology (PEM) developed by the Identity Ecosystem Steering Group (IDESG) public-private partnership working to implement the NSTIC, and NIST Special Publication (SP) 800-53 Rev. 4 Appendix J, which contains the Privacy Control Catalog. FIPPs provide the basis for virtually every sector-specific federal privacy law and regulation—including the Health Insurance Portability and Accountability Act (HIPAA) governing healthcare and the Gramm-Leach-Bliley Act (GLBA) statute governing financial services—and for numerous state privacy laws and regulations as well. A set of FIPPs has also been promulgated and its use advocated by the Federal Trade Commission for the electronic marketplace.

The FIPPs serve as a useful guidepost for generally addressing the foundational notions of privacy. They are value statements rather than recipes, however. System planners often encounter difficulties when trying to operationalize them, particularly when assessing privacy risks and when establishing privacy requirements for designing and developing systems and technologies. Those activities require a well-articulated set of privacy objectives and a privacy risk assessment approach, from which privacy risks can be evaluated and implementation requirements can be developed. An example of these difficulties is discussed in the paragraphs that follow, with terminology aligned with the definitions from the NSTIC FIPPS provided in Table 1.

Table 1. NSTIC FIPPs

NSTIC FIPPs
Transparency: Organizations should be transparent and notify individuals regarding collection, use, dissemination, and maintenance of personally identifiable information (PII).
Individual Participation: Organizations should involve the individual in the process of using PII and, to the extent practicable, seek individual consent for the collection, use, dissemination, and maintenance of PII Organizations should also provide mechanisms for appropriate access, correction, and redress regarding use of PII.
Purpose Specification: Organizations should specifically articulate the authority that permits the collection of PII and specifically articulate the purpose or purposes for which the PII is intended to be used.
Data Minimization: Organizations should only collect PII that is directly relevant and necessary to accomplish the specified purpose(s) and only retain PII for as long as is necessary to fulfill the specified purpose(s).
Use Limitation: Organizations should use PII solely for the purpose(s) specified in the notice Sharing PII should be for a purpose compatible with the purpose for which the PII was collected.
Data Quality and Integrity: Organizations should, to the extent practicable, ensure that PII is accurate, relevant, timely, and complete
Security: Organizations should protect PII (in all media) through appropriate security safeguards against risks such as loss, unauthorized access or use, destruction, modification, or unintended or inappropriate disclosure
Accountability and Auditing: Organizations should be accountable for complying with these principles, providing training to all employees and contractors who use PII, and auditing the actual use of PII to demonstrate compliance with these principles and all applicable privacy protection requirements

Risk is typically expressed as a function of threats and vulnerabilities as well as the likelihood of those materializing. While it is possible to identify threats and vulnerabilities for an organization exhibiting certain values, in this case the FIPPs, the ability to assess risk

[9] http://www.whitehouse.gov/sites/default/files/rss_viewer/NSTICstrategy_041511.pdf

[10] https://www.privacyassociation.org/resource_center/privacy_glossary#F, accessed January 27, 2014.

fails in two key ways: 1) individual perception of whether and how an organization exhibits the values is subjective, and 2) compliance doesn't address the full scope of privacy risks. In this case of *Transparency*, for example, a threat may be that the mechanism for providing notification, such as providing a website privacy policy, becomes unavailable. An associated vulnerability may be that the web server that serves the privacy policy is sitting on a single physical server in a data center with an unreliable power supply. A reasonable measure of likelihood for web server up-time in light of the unreliable power supply could probably be determined, but the likelihood of an individual feeling like they have received adequate notice cannot be easily or accurately assessed. To some, the act of providing notification is enough to claim compliance with the Transparency FIPP. This means that in the case of the web server availability, there may be an occasional compliance gap to the organization, but no other risks may be identified. This ignores the risks to the individual. If the notice is not effective, the value of transparency is not met. This leaves organizations and individuals potentially open to privacy risk. Questions such as the following must be answered to better understand whether risks around the particular FIPP were addressed:

- Is the notice accurate and effective?
- How would an organization determine the likelihood that a notification is provided in a way that the individual interprets it correctly?
- Does an organization know whether each individual that is required to receive the notice has in fact received it?
- How does the organization ensure that the activities of its systems are consistent with what is described in the notice?

The 3rd and 4th bullets above can be supported by technical capabilities of the system. Taking a privacy engineering approach will help ensure that happens.

Another challenge with using the FIPPs as a "risk tool" is that they are articulated in such broad terms that the do not have enough specificity to govern a particular system based on the context of how it collects and uses PII. *Security* is one example of this. The Security principle is generically discussed in terms of safeguarding information from certain risks, but the scope of what "safeguarding" truly entails is not addressed or even implied. Based on current work in the National Security community to identify the security controls necessary for safeguarding PII in information systems, over 100 controls are required to meet the requirements for government agencies that are defined in applicable laws and OMB policies[11]. Another example is *Data Quality and Integrity*, which requires ensuring that PII is "accurate, relevant, timely, and complete". This can vary widely depending on the purpose for collecting PII. The same PII data set could have radically different requirements from one system to another.

While privacy poses particular challenges in accounting for context (and some noteworthy approaches, such as contextual integrity[12], have arisen in response), it would be fruitful to evolve existing systems engineering techniques to help address this. One such technique,

[11] The Committee on National Security Systems (CNSS) is developing the Privacy Overlay, which will be an attachment to Committee on National Security Systems Instruction (CNSSI) No. 1253, *Security Categorization and Control Selection for National Security Systems*, Appendix F, *Overlays*

[12] H. Nissenbaum, Privacy in Context: Technology, Policy, and the Integrity of Social Life, Stanford, CA: Stanford University Press, 2010.

familiar to all systems engineers, is tradespace analysis. Tradespace analysis represents explicit acknowledgement that developing practical systems invariably involves trading off some capabilities and properties against one another. In a world of finite resources, the classic tradeoff calculations remain capabilities versus cost, capabilities versus schedule and cost versus schedule, but developing a system of any complexity will also involve many other tradespaces. Privacy further broadens this reality in the form of potential tradeoffs between not only privacy and other attributes, but also between different aspects of privacy. A frequently observed example of the former is the tradeoff between the strength of anonymization and the utility of the anonymized data.

A less obvious example of the latter is the tradeoff between the linkability afforded by different biometric modalities and the ancillary personal information that may be derived from them.

Systematically identifying such tradeoffs would constitute an advance, but real dividends would come from techniques for rigorously analyzing them in support of the necessary engineering judgments. Existing techniques from systems engineering might be adapted to address tradeoffs involving privacy, but privacy may also require methods that are specific to it[13]. While good engineering should minimize tradespaces, it cannot eliminate them altogether. A key objective for tradespaces and tradespace analyses involving privacy should be techniques that appropriately minimize the tradespaces and that enable analysis supporting appropriate valuations of privacy capabilities and properties from the standpoint of society as well as individuals.

As discussed under Question 2, the commonly referenced FIPPs alone are not enough to articulate how privacy risks should be managed for systems. To achieve privacy objectives, privacy requirements must be built-in with the same rigor as requirements for system functionality. Various engineering models refer to requirements in certain ways, such as distinguishing between "functional" and "non-functional" requirements. How privacy requirements are typified is not as important as ensuring that they are included in system design considerations and that they are described at a level that is usable by system engineers. Taking *Use Limitation* as an example, if an engineer is provided a requirement that states, "The system must use PII solely for the purpose(s) specified in the notice," they do not have adequate information to design and develop a system that meets everything intended by those requirements. Program owners must think about what that means in the context of their system and provide a more specific statement that is tailored for their system. For example, consider a system that generates correspondence to mail to individuals. The system only needs enough PII to mail the letter, not an individual's entire record. If a developer is provided requirement statements that read, "PII input from users shall be limited to the following data elements: Name, account number, and mailing address," they will know how to design the user interface, how to design the database, and how to design system interfaces, all in a way that controls the use of the system in a way that is consistent with its defined purpose of generating and mailing correspondence.

[13] It should be noted that the results of such analysis could potentially contravene the "Full Functionality – Positive-Sum, not Zero Sum" Privacy by Design Foundational Principle (see http://www.privacybydesign.ca/index.php/about-PbD/7-foundational-principles). However, this principle is most properly viewed as philosophical rather than an engineering dictum.

There are three primary sources of deriving privacy requirements: 1) principles, which are typically found in applicable laws, regulations, directives, and policies, 2) domain-specific risks, which are influenced by things like industry, organizational mission, and compliance requirements, and 3) the purpose and functionality of the system. The first two comprise baseline requirements and the third comprises risk-based requirements. In all cases, controls should be selected and implemented with cognizance of potential failure modes related to privacy. All fields of engineering aim to avoid single points of failure in which the failure of one component or control results in system failure. Unfortunately, this concern has been largely absent when attempts are made to ensure privacy within socio-technical systems[14].

Question 3: Multi-disciplinary approach

Discuss how privacy challenges and objectives might be framed to bring many disciplines (e.g., computer science, economics, social and behavioral sciences, and law disciplines) together to jointly and collaboratively work to both strengthen privacy and support innovation in cyberspace and information systems; discuss how diverse national/cultural perspectives on privacy can be accommodated.

After decades of research and practical experience, it has become quite clear that privacy resides at the intersection of numerous disciplines ranging across the humanities, social sciences, law, business, engineering, and the hard sciences. Therefore, any effort to develop a coherent and useful multi-disciplinary approach must cast a wide net. Given that much of the consternation regarding privacy over the last half century has been driven by technical developments within social contexts, it is particularly important not to overlook the potential application of work in Science and Technology Studies (STS). An inter-disciplinary field in its own right, STS consists of a select set of techniques from the humanities and social sciences applied to understanding the development and operation of socio-technical systems. These techniques include historical synthesis[15], ethnomethodology[16], and actor-network theory[17]. In various different ways at various different levels, all these techniques concern themselves with the construction of meaning and, as such, have direct application to understanding how those meanings associated with privacy are enabled, undermined, altered, and otherwise manifested by systems. Therefore, STS may prove a rich source of methods for supporting the analysis and embedding of privacy in socio-technical systems. A National Privacy Research Strategy should include adaptation and piloting of relevant STS methods for use by engineers and policy makers, among others.

By the same token, work on the governance of human subjects research may also prove relevant for the Strategy. The potential salience of approaches to human subjects research

[14] The use of anonymization in publicly released data sets has become a classic instance of this. Assessments of privacy must include systematic analysis of potential failure modes so as to achieve control coverage sufficient to minimize or eliminate single points of failure.
[15] For example, see W.G. Vincenti, *What Engineers Know and How They Know It: Analytical Studies from Aeronautical History*, Baltimore, MD: Johns Hopkins University Press, 1990.
[16] For example, see E. Wegner, *Communities of Practice: Learning, Meaning, and Identity*, Cambridge: Cambridge University Press, 1998.
[17] For example, see B. Latour, *Reassembling the Social: An Introduction to Actor-Network Theory*, New York: Oxford University Press, 2005.

protection for "Big Data" and privacy has already been pointed out[18], however, with appropriate development, the potential applicability may be broader. In particular, approaches to risk calculation and management that are part of current protection requirements[19] could be expanded upon to provide a more rigorous basis for thinking about privacy risk. This expansion could include work on risk-risk tradeoffs (relevant for tradeoff analysis) as well as population risk and benefit mapping. The importance of such processes for privacy have been brought into high visibility recently by the widely criticized experiments that Facebook and OkCupid have performed on their users.

Privacy by Design (PbD) advances the view that privacy cannot be assured solely by compliance with regulatory frameworks; rather, privacy assurance must become an organization's default mode of operation. PbD applies to information technology, accountable business practices, and physical design. Simply stated, privacy is not ensured by policy alone. Adequate privacy requires thoughtful integration with every layer of an organization, including:

- Organization policies and governance;
- Business processes;
- Standard operating procedures;
- System and network architectures;
- IT system design and development practices
- Management of data sources

Privacy engineering is a systematic, risk-driven process that operationalizes the Privacy by Design philosophical framework within IT systems by:

- Segmenting PbD into activities aligned with those of the systems engineering life cycle (SELC) and supported by particular methods that account for privacy's distinctive characteristics
- Defining and implementing requirements for addressing privacy risks within the SELC using architectural, technical point, and policy controls. Privacy requirements must be defined in terms of implementable system functionality and properties. Privacy risks, including those beyond compliance risks, are identified and adequately addressed.
- Supporting deployed systems by aligning system usage and enhancement with a broader privacy program

The goal is to integrate privacy into existing systems engineering processes; it is not to create a separate new process.[20]

Figure 1 illustrates how the core privacy engineering activities map to stages of the classic systems engineering life cycle. A mapping exists for every systems engineering life cycle, including agile development, since every life cycle includes the core activities in some form.

---

[18] For example, see R. Calo, "Consumer Subject Review Boards: A Thought Experiment," *Stanford Law Review Online,* **66**: p. 97-102, 2013.

[19] http://www.hhs.gov/ohrp/humansubjects/commonrule/

[20] MITRE Corporation, *The MITRE Systems Engineering Guide: Privacy Engineering*, http://www.mitre.org/publications/systems-engineering-guide/systems-engineering-guide

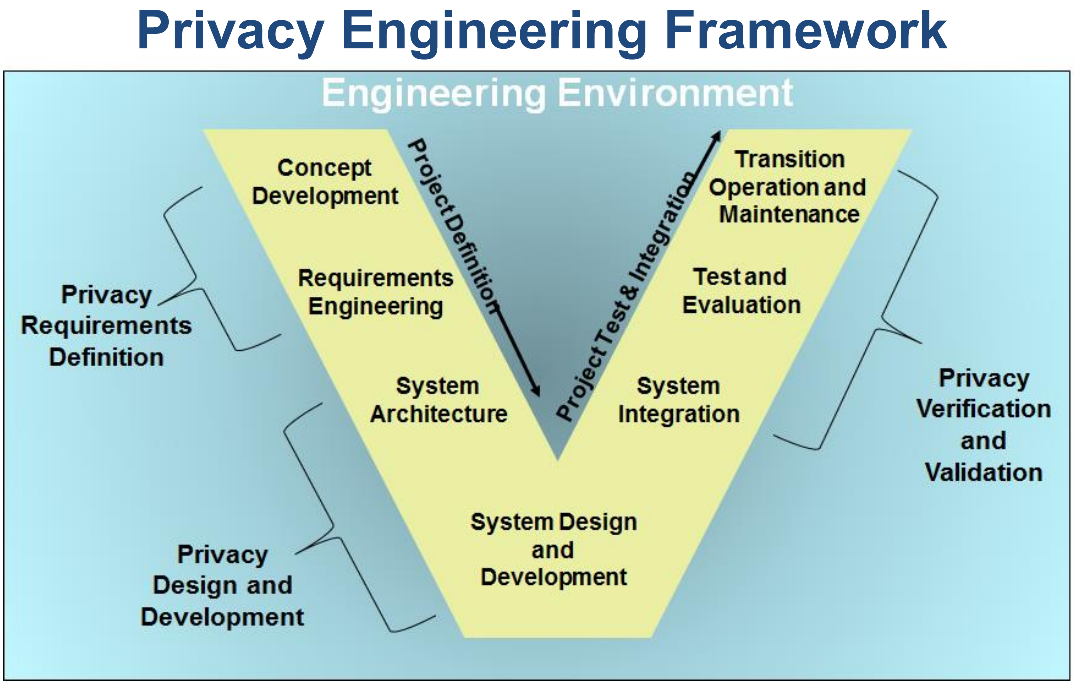

Figure 1. Privacy Engineering Framework[21]

[21] MITRE Corporation, *Privacy Engineering Framework*, http://www.mitre.org/privacy

The primary privacy engineering activities and methods are listed in the Table 2 below.

Table 2. Privacy Engineering Activities and Methods[22]

Life Cycle Activity	Privacy Method	Method Description
Privacy Requirements Definition: Specification of system privacy properties in a way that supports system design and development	Baseline & custom privacy system requirements	Granular technical privacy requirements derived from first principles and from risk analysis
	Privacy empirical theories & abstract concepts	Methodological constructs based on theories of privacy and socio-technical systems
Privacy Design and Development: Representation and implementation of those elements of the system that support defined privacy requirements	Fundamental privacy design concepts	Explicit or tacit consensus understandings of how privacy works in a system
	Privacy empirical theories and abstract concepts	Methodological constructs based on theories of privacy and socio-technical systems
	Privacy design tools	Specific techniques for achieving privacy
	Privacy heuristics	Experientially developed rules of thumb regarding privacy properties of artifacts
Privacy Verification and Validation: Confirmation that defined privacy requirements have been correctly implemented and reflect stakeholder expectations	Privacy testing & review	Executable tests and targeted document reviews associated with privacy requirements
	Operational synchronization	Analysis of privacy policies & procedures and system behaviors for inconsistencies

Question 4: Privacy architectures

(a) The Big Data report recommends adoption of a "responsible use framework" that would provide greater focus on the use of data and hold entities that utilize data accountable for responsible use of the data. Describe an architecture implementing a "responsible use framework" incorporating the three questions above and taking into account issues as: Encoding privacy policies in machine-checkable forms and ensuring their compliance and auditability; managing the collection, retention, and dissemination of sensitive data; and ensuring the confidentiality and integrity of sensitive data; while enabling desired uses of them. (b) Describe other privacy architectures that would be effective for the design and implementation of privacy-preserving information systems. (c) Describe technological advances that can change privacy perceptions and how those advances would be incorporated into the "responsible use framework" architecture or other architectures submitted for 4(b).

The concept of responsible use oversimplifies the realities of grounded practice with respect to the development and operation of socio-technical systems. For example, much has been made of research results demonstrating the detection of an adverse drug interaction using Internet search data. The conclusion to which many Big Data proponents immediately leapt was the practical impossibility of applying collection minimization and use limitation to such projects. When one examines what took place, though, a considerably different picture emerges.[23] The real ground-breaking research, rather than taking place on Internet search data, took place on data in the FDA Adverse Event Reporting System

[22] MITRE Corporation, *Privacy Engineering Framework*, http://www.mitre.org/privacy

[23] N.P. Tatonetti, G.H. Fernald, and R.B. Altman, "A novel signal detection algorithm for identifying hidden drug-drug interactions in adverse event reports," *J Am Med Inform Assoc,* **19**(1): p. 79-85, 2012; R.W. White et al., "Web-scale pharmacovigilance: listening to signals from the crowd," *J Am Med Inform Assoc,* **20**(3): p. 404-8, 2013.

(FAERS). Novel data mining algorithms applied to the data in this system identified several previously unrecognized adverse drug interactions. The researchers then attempted to detect one of those interactions using historical Internet search data. This required the development of a set of terms related to the symptoms associated with the interaction in the FAERS data. The researchers successfully found in the Internet search data the interaction they were specifically looking for using the symptoms they had already identified.

The actualities of these kinds of studies militate against an uncritical emphasis on responsible use. In those described above, the high-impact research was performed on a database that existed specifically for the purpose for which it was being used. The Internet results depended on that usage, and the historical search information used in obtaining those results had been collected and used with the permission of the individuals using the search engine. Responsible use may merit a place in the Strategy, but so do those aspects of privacy practice that responsible use is supposedly rendering passé. Privacy architectures must be holistic rather than focused on responsible use. The realities of grounded practice continue to offer opportunities for effective privacy controls at every stage of the information life cycle.

One of the tenets of the Privacy by Design movement is that privacy assurance becomes a default mode of operation. This points to a need to include privacy in an organization's enterprise architecture. The natural progression would be to incorporate privacy into software architecture. Managing privacy risks through a well-defined architecture is always preferable, but reality dictates that privacy controls in any system will consist of a variety of architectural, point, and policy controls. To be effective, privacy controls must not only be considered in the context of a specific system, but also in the context of all system activities that influence how systems are built. Figure 2 below represents the functional layers of system activities.

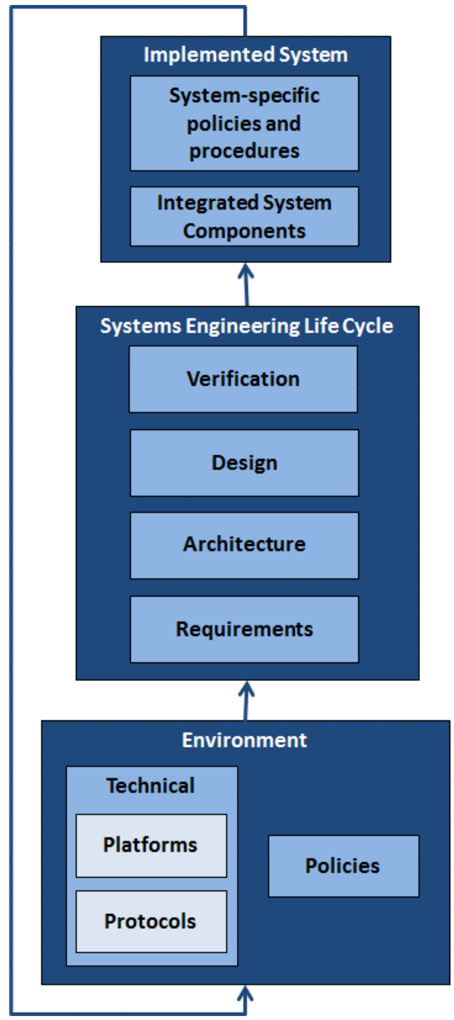

Figure 2. System Activities Stack

Privacy must be addressed in each layer of engineering and operational activities. This ensures privacy is an organic part of the system development and deployment processes that is consistently addressed throughout all activities for all systems, rather than an unnatural "bolt on" activity singularly addressed by system owners.

Similarly, consideration of privacy-enhancing technologies (PETs), let alone application of PETs, is hobbled by the lack of a sufficiently rich generally-accepted taxonomy for discussing and thinking about them. While a number of relatively granular taxonomies have been proposed[24], none have gained any traction. Coarser terminology[25] has not fared much better. The situation is further complicated by PETs that demand altered worldviews, such as differential privacy[26], an approach that treats privacy as a resource that must be sustainably managed and that is depleted by each use of the technology.

[24] For example, see C. Adams, "A classification for privacy techniques," *University of Ottawa Law and Technology Journal*, 3(1):35–52, 2006.

[25] For example, see I. Rubinstein, "Regulating Privacy by Design," *Berkeley Technology Law Journal,* **26**: p. 1409, 2012.

[26] C. Dwork, "A Firm Foundation for Private Data Analysis," *Communications of the ACM,* **54**(1): p. 86-95, 2011.

What is needed is both wider and deeper than anything yet proposed. More than a typology, this common vocabulary must also encompass and support selection of architectural, discrete technical, and policy privacy controls at appropriate points in the systems activities stack (see Figure 2). Further, it must also address various quality attributes—performance, usability, etc.—in terms that are specific to PETs. Indeed, it can be argued that, generally speaking, quality attributes constitute the critical path for PET evolution and adoption.

For example, a number of different PETs rely in some way on secure multi-party computation (SMPC). SMPC, though, is currently computationally expensive and therefore impractical in most settings. Reducing the time complexity of SMPC is therefore crucial. SMPC-based PETs are also not the most user friendly of technologies for non-specialist system developers, system operators, and system users. While continued research into novel PETs is certainly worth pursuing, the reverse salient[27] impeding the further development and adoption of many extant PETs is quality attributes.

Conclusion and Recommendations

Information technology is a permanent fixture in our society and economy. It represents both one of our greatest sources of innovation and some of the most challenging privacy issues faced to date. It also represents a wonderful and compelling opportunity to merge the best of innovation with the best of privacy practices so that we may continue to enjoy the benefits of each. Research in this area is much needed and well-timed. There is a growing body of research in many areas of privacy that can inform the National Privacy Research Strategy. Focusing the strategy on PETs will result in important progress on the side of privacy. Perhaps even more important is focusing on the methods required to design and develop privacy-aware systems and technologies. MITRE recommends the Strategy incorporate the following considerations as discussed throughout this response:

- Aim for solutions that require minimal effort by individuals to maintain their desired level of privacy,
- Identify opportunities to automate privacy controls such that privacy is an organic part of operations for systems and their users,
- Acknowledge the importance of the FIPPs, but look beyond them to other privacy risks and harms that the FIPPs do not and cannot inherently address,
- Look within engineering disciplines and also beyond,
- Define methods that are systematic, risk-driven, repeatable, and a natural fit within the systems engineering life cycle,
- Explore the related work to be done in all of the layers of the Systems Activities Stack (Figure 2), and
- Acknowledge that although paradigm shifts must occur, the path forward does not have to be radical or complicated.

[27] An explanation of the concept of a reverse salient can be found in T.P. Hughes, *Networks of Power: Electrification in Western Society, 1880-1930*, reprint, Baltimore, MD: Johns Hopkins University Press, 1993.

Appendix A: Examples of FIPPs Frameworks

- The Organization for the Advancement of Structured Information Standards (OASIS) has an active Privacy Management Reference Model (PMRM) Technical Committee. Their 2012 work product is available here: http://docs.oasis-open.org/pmrm/PMRM/v1.0/csd01/PMRM-v1.0-csd01.pdf
- Similar to the present initiative is the work of the Identity Ecosystem Steering Group (IDESG) stood up in August of 2012 as a private sector led initiative in response to the White House's National Strategy for Trusted Identities in Cyberspace (NSTIC). The IDESG Privacy Coordination Standing Committee has created a Privacy Evaluation Methodology (PEM) that is available here: http://www.idecosystem.org/group/privacy-coordination-committee
- Also similar to the present initiative is the work done by the Smart Grid Interoperability Panel. Through the support of a joint Federal-private sector group, NIST published NIST Information Report (NISTIR) 7628 with its own separate volume specifically on the issue of privacy and the Smart Grid. It is available here: http://csrc.nist.gov/publications/nistir/ir7628/nistir-7628_vol2.pdf
- The Organization for Economic Cooperation and Development (OECD) has been working on Privacy and Fair Information Practice Principles for several decades. Its 1980 work is listed here: http://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm
- OECD's more recent Privacy Framework is available here: http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf
- The Asia Pacific Economic Cooperation published its Privacy Framework in 2005. It has subsequently addressed privacy issues in related publications. The Privacy Framework is available here: http://publications.apec.org/publication-detail.php?pub_id=390